

Analysis of stability of community structure across multiple hierarchical levels

HUI-JIA LI¹ and XIANG-SUN ZHANG^{(a)2,3}

¹ School of Management Science and Engineering, Central University of Finance and Economics, Beijing 100080, China.

² Academy of Mathematic and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.

³ National Center for Mathematics and Interdisciplinary Sciences, Chinese Academy of Sciences, Beijing 100190, China.

PACS 89.75.Hc – First pacs description
 PACS 89.75.Fb – Second pacs description

Abstract – The analysis of stability of community structure is an important problem for scientists from many fields. Here, we propose a new framework to reveal hidden properties of community structure by quantitatively analyzing the dynamics of Potts model. Specifically we model the Potts procedure of community structure detection by a Markov process, which has a clear mathematical explanation. Critical topological information regarding to multivariate spin configuration could also be inferred from the spectral significance of the Markov process. We test our framework on some example networks and find it doesn't have resolute limitation problem at all. Results have shown the model we proposed is able to uncover hierarchical structure in different scales effectively and efficiently.

1. Introduction. – Community structure detection [1, 2] is a main focus of complex network studies. It has attracted a great deal of attentions from various scientific fields. Intuitively, community refers to a group of nodes in the network that are more densely connected internally than with the rest of the network. A well known exploration for this problem is the concept of modularity, which is proposed by Newman et al [1, 2] to quantify a network's partition. Optimizing modularity is effective for community structure detection and has been widely used in many real networks. However, as pointed out by Fortunato et al [3], modularity suffers from the resolution limit problem which is concerned about the reliability of the communities detected through the optimization of modularity. Complementary to the modularity concept, many efforts are devoted to understanding the properties of the dynamical processes taking place in the underlying networks. Specifically, researchers have begun to investigate the correlation between the community structure and the dynamical systems, such as synchronization [4] [5] and random walk process [6] [13].

Potts dynamical model is a powerful tool which has been applied to uncover the thermodynamical behaviors in networks [7, 8]. It models an inhomogeneous ferromagnetic system where each node is viewed as a labeled spin in the network. The configuration of the system is defined by the interactions between the nodes. Considering an unweighted network with N nodes without self-loops, a spin configuration $\{S\}$ is defined by assigning each node i a spin label s_i which may take integer values $s_i = 1, \dots, K$. To characterize the coherence between two nodes, spin-spin correlation $C = (C_{ij})$ is defined as the thermal average of δ_{s_i, s_j} :

$$C_{ij} = \langle \delta_{s_i, s_j} \rangle, \quad (1)$$

which represents the probability that spin variables s_i and s_j have the same value. C_{ij} takes values from the interval $[0, 1]$, representing the continuum from no coupling to perfect accordance of nodes i and j . In section 4, we develop a novel hierarchical block model which can calculate C efficiently. The C is corresponding to the average spin correlation across multiple level of the hierarchical structure.

If the system is not homogeneous but has a community

^(a)Corresponding authors: zxs@amt.ac.cn

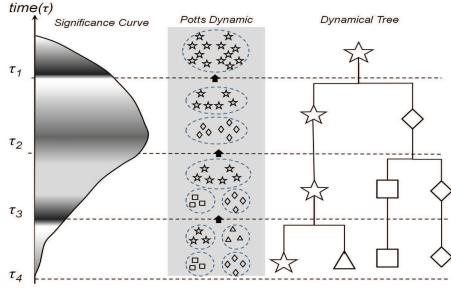


Fig. 1: Dynamics of spin configuration of four communities when they go through several local uniform states to the global stable state. Different spin values are described by different shapes in dynamical tree. τ_i represents the time at which the system has i different spin states

. The significance curve of spin configurations at different times can be calculated shown in the left part of the graph.

structure, the states are not just ferromagnetic or paramagnetic [7]. We assume that the spins will go through a hierarchy of local uniform states (meta-stable states) as time increases which is shown in Fig.1, before they reach a globally stable state with the same value. In each local uniform state, spin values of nodes within the same communities are identical. Correspondingly, one can calculate the hitting and exiting time of each local uniform state and there should be a big gap between them when a well-formed community structure exists. The significance of spin configurations at different time can also be calculated to illustrate the amplitude of variation.

In this letter, using the Potts model and spectral theory, we firstly uncover the relationship between community structure of a network and its meta-stability of spin dynamics, and then propose the significance of communities to characterize and analyze the underlying spin configuration. For any given network, one can straightforwardly get critical information related to its community structure, such as the stability and the optimal number of communities across multiple timescales without using particular partition algorithms. We then use phase transition of stochastic dynamical system to prove that the stability we proposed is able to indicate the significance of community structure more theoretically which is based on eigengap theory. Furthermore, a novel hierarchical block model is proposed which can calculate spin correlation at each layer of the network structure. Finally, we test our framework on some examples of complex networks. Results show the model we proposed is able to uncover the hierarchical structure in different scales effectively and efficiently and doesn't have resolute limitation problem at all.

2. The framework. – In order to establish the connection between the community structure and the local uniform behavior of Potts model, we introduce a Markov stochastic model featured by spectral significance for the

network. Markov process is a useful tool and has been applied to find communities [6]. Let $P = (p_{ij})$ be the stochastic transition matrix and the element p_{ij} is defined as

$$p_{ij} = \frac{C_{ij}}{\sum_{j=1}^N C_{ij}} \quad (2)$$

where C_{ij} is the spin-spin correlation function defined in Eq.(1). Via this representation, the tools of stochastic theory and finite-state Markov processes [6] can be utilized for the purpose of community structure analysis.

For this ergodic Markov process, P^t represents the transition probability matrix between nodes over a period of t time steps. To compute the transition matrix P^t , the eigenvalue decomposition of P is used. If λ_k with $k = 1, \dots, n$ denote the eigenvalues of P , and its right and left eigenvectors u_k and f_k are scaled to satisfy

$$Pu_k = \lambda_k u_k, f_k P = \lambda_k f_k \quad (3)$$

The orthonormality relation of u_k and f_l is satisfied:

$$u_k f_l = \delta_{kl}, \quad (4)$$

and the spectral representation of P is given by

$$P^t = \sum_k \lambda_k^t u_k f_k \quad (5)$$

We assume that eigenvalues of P are sorted such that $\lambda_1 = 1 > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$. Because of the largest eigenvalue $\lambda_1 = 1$, when time $t \rightarrow \infty$, $P^{(0)} = P^\infty = u_1 f_1$. The convergence of every initial distribution to the stationary distribution $P^{(0)}$ corresponds to the fact that the spins of whole system ultimately reach exactly the same value, as time increases. This perspective belongs to a timescale $t \rightarrow \infty$, at which all eigenvalues λ_k^t go to 0 except for the largest one, $\lambda_1^t = 1$. In the other extreme of a timescale $t = 0$, P^t becomes the stationary distribution matrix. All of its columns are different, and the system disintegrates into as many spin values as possible. Then, we simply extend P^t to the symmetrical form $G^{(t)} = (P^t + (P^t)^T)/2$.

Suppose the partition method divides the network A into K communities or sets $V_k \subset V, k \in \{1, 2, \dots, K\}$ which are disjoint and the sets V_1, V_2, \dots, V_K together form a partition of node set V . The number of nodes in each community is denoted by $N_k = |V_k|$. We take the time series into consideration. Therefore, we define the significance of a given community k by the ratio of inner correlations as

$$S_k^{(t)} = \sum_{i,j \in V_k} \frac{[G^{(t)}]_{i,j}}{N_k} \quad (6)$$

$S_k^{(t)}$ can be viewed as a function of timescale t and we can use it to study the trend of community structure as time goes on.

Further discussion is facilitated by reformulating the average association objective in matrix form. We denote the membership vector of community k by x_k , a binary vector that describes each node's involvement in community k . The hard partition and disjointness of sets V_k requires that the vectors x_l and x_s are orthogonal. Given the number of communities K , the communities are found by maximizing the objective function

$$J_K^{(t)} = \sum_{k=1}^K \sum_{i,j \in V_k} \frac{[G^{(t)}]_{i,j}}{N_k} = \sum_{k=1}^K \frac{x_k^T G^{(t)} x_k}{x_k^T x_k} \quad (7)$$

The objective is to be maximized under the conditions $x_k \in \{0, 1\}$ and $x_l^T x_s = 0$ if $l \neq s$. Eq.(7) can be rewritten as a matrix trace by accumulating the vectors u_k into a matrix $X = (x_1, x_2, \dots, x_K)$. We can then write the objective $J_K^{(t)}$ as

$$\begin{aligned} J_K^{(t)} &= \text{tr}\{(X^T X)^{-1} X^T G^{(t)} X\} \\ &= \text{tr}\{(X^T X)^{-1/2} X^T G^{(t)} X (X^T X)^{-1/2}\} \end{aligned} \quad (8)$$

where matrix $X^T X$ is diagonal. The substitution $Y = X(X^T X)^{-1/2}$ simplifies the optimization problem to $J_K^{(t)} = \text{tr}\{Y^T G^{(t)} Y\}$. The condition $Y^T Y = I_K$ is automatically satisfied since

$$Y^T Y = (X^T X)^{-1/2} (X^T X) (X^T X)^{-1/2} = I_K. \quad (9)$$

The vectors y_k thus have unit length and are orthogonal to each other. The optimization problem can be written in terms of the matrix Y as

$$\max_{Y^T Y = I} \text{tr}\{Y^T G^{(t)} Y\}. \quad (10)$$

According to Rayleigh-Ritz theorem [14], the maximum for this problem is attained when columns of Y is the right eigenvectors $U = \{u_1, \dots, u_K\}$ corresponding to the K largest eigenvalues of the symmetric correlation matrix $G^{(t)}$. Then the strength of such a community is approximately equal to its corresponding t -th power of the eigenvalue

$$S_k^{(t)} \approx \frac{u_k^T G^{(t)} u_k}{u_k^T u_k} = \lambda_k^t \frac{u_k^T u_k}{u_k^T u_k} = \lambda_k^t \quad (11)$$

For the convergence of the Potts model across multiple timescales, the vanishing of the smaller eigenvalues as the time growing describes the loss of different spin states and the removal of the structural features encoded in the corresponding weaker eigenvectors. For the purpose of community identification, intermediate timescales of local uniform states are interesting. If we want to identify z communities, we expect to find P^t at a given timescale, the eigenvalues λ_k^t may be significantly different from zero only for the range $k = 1, \dots, z$. This is achieved by determining t such that $|\lambda_k|^t \approx 0$.

From another perspective, because the eigenvalues are sorted by $\lambda_1 = 1 > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$, the strength of a community at time t , λ_k^t , can also be viewed as the robustness of k -spin state at time t . At this point, the eigengap $\lambda_{k-1}^t - \lambda_k^t$ can be interpreted as the “difficulty” that the k -spin state transfer to the $(k-1)$ -spin state at time t . The number of communities Λ at time t is then inferred from the location of the maximal eigengap, and this maximal value can be used as a quality measure for the most stable state. The $\Lambda(t)$ is formally defined as

$$\Lambda(t) = \arg[\max_k (\lambda_{k-1}^t - \lambda_k^t)] \quad (12)$$

From a global perspective if the number of communities Λ doesn't change for the longest time, we can consider it as the optimal number for this network, represented as Ψ .

To a certain extent, the most stable state can represent the spin configuration of the whole network. Thus, we define the stability of community structure at each timescale, $\Theta(t)$, as the stability of the most stable spin state:

$$\Theta(t) = \lambda_{\Lambda(t)-1}^t - \lambda_{\Lambda(t)}^t \quad (13)$$

Our expectation is that from the trend of $\Theta(t)$, one can find the most stable timescale for community structure where $\Theta(t)$ reaches the maximal. Furthermore, from a global perspective, we can use the largest stability corresponding to q communities, $\Gamma(q) = \max\{\Theta(t) | \Lambda(t) = q\}$, to indicate the robustness of a network, defined as the stability of the structure with q communities. While $\Gamma(q)$ tries to directly characterize the network structure rather than a specific network partition thus very convenient to estimate the modularity property of the network.

3. Prove the validity of stability. — Many measures have been defined to indicate the significance of community structure, such as modularity Q proposed by Newman et al [1] [11] and spectral cut metrics [14]. In [10], the eigenvector of transition matrix P is also found able to indicate the partition of the nodes in the network. The components of eigenvector corresponding to nodes within the same community have very similar values and the eigenvalue gaps between different communities can represent the significance of the modularity structure. In this part, we use phase transition of stochastic dynamic system to prove that the stability we proposed is in proportion to the eigenvalue gap. Thus, the larger the stability, the larger the significance of the community structure.

Let us demonstrate our argument for the simplest case that a network owning 2 communities. As $\lambda_1 \equiv 1$, according to Eq.(5), we write

$$P^t = u_1 f_1 + \lambda_2^t u_2 f_2 + |\lambda_3|^t B^t, \quad (14)$$

where B^t is the remainder matrix

$$B^t = \sum_{k \geq 3} \frac{\lambda_k^t}{|\lambda_3|^t} u_k f_k \quad (15)$$

Although P need not to be diagonalizable, the representation is guaranteed by the nondegeneracy of λ_1 and λ_2 . Because $\sum_x f_2(x)u_1(x) = 0$ and $u_1 > 0$, we deduce that

$$f_M \equiv \max_{x \in X} f_2(x) > 0 > f_m \equiv \min_{x \in X} f_2(x) \quad (16)$$

Let us now fix $a > 0$ and according to Eq.(16), we consider the two nodes sets belonging to two different communities:

$$\begin{aligned} I_M(a) &= \{x \in X \mid \frac{f_M - f_2(x)}{f_M} < a\}, \\ I_m(a) &= \{x \in X \mid \frac{f_m - f_2(x)}{f_m} < a\}, \end{aligned} \quad (17)$$

These two sets will in fact turn out to be the two phases. We also define $E_M(a) \equiv \varphi I_M(a)$, $E_m(a) \equiv \varphi I_m(a)$ (where φ indicates the complement of a set). We take $a < 1$ and note that

$$a < 1 \Rightarrow I_M(a) \text{ and } I_m(a) \text{ are disjointed.} \quad (18)$$

Accordingly to the two nodes sets, the corresponding the two phases of the system are,

$$\begin{aligned} u_M^t(x) &= u_1(x) + \lambda_2^t u_1(x) f_M + |\lambda_3^t| B_{xM}^t, \\ u_m^t(x) &= u_1(x) + \lambda_2^t u_1(x) f_m + |\lambda_3^t| B_{xm}^t, \end{aligned} \quad (19)$$

Here B_{xM}^t (respectively B_{xm}^t) is the value of B_{xy}^t for some point y ($\equiv y_M$) (respectively, y_m) such that $f_2(y) = f_M$ (respectively, f_m). From Eq.(19), we have

$$\begin{aligned} u_1(x) &= \frac{f_M u_m^t(x) - f_m u_M^t(x)}{f_M - f_m} + |\lambda_3^t| \frac{f_m B_{xM}^t - f_M B_{xm}^t}{f_M - f_m} \\ u_2(x) &= \frac{f_M^t(x) - u_m^t(x)}{\lambda_2^t(f_M - f_m)} + \frac{|\lambda_3^t| B_{xm}^t - B_{xM}^t}{f_M - f_m} \end{aligned} \quad (20)$$

More generally, for any y one can define

$$u_y^t(x) = u_1(x) + \lambda_2^t u_2(x) f_2(y) + |\lambda_3^t| B_{xy}^t \equiv P_{xy}^t \quad (21)$$

We take the scalar product of u_M^t and u_m^t with f_2 . There are

$$\sum_x f_1(x) B_{xy}^t = 0. \quad (22)$$

This follows from

$$B^t = \sum_{k \geq 3} \frac{\lambda_k^t}{|\lambda_3^t|} u_k f_k \quad (23)$$

and therefore

$$\sum_x f_2(x) u_M^t(x) = \lambda_2^t f_M, \quad \sum_x f_2(x) u_m^t(x) = \lambda_2^t f_m, \quad (24)$$

From Eq.(24) and the fact that $\sum u_M^t(x) = \sum u_m^t(x) = 1$ we can deduce the basic identity

$$\lambda_1^t - \lambda_2^t = \sum_x u_M^t(x) \left(1 - \frac{f_2(x)}{f_M}\right) = \sum_x u_M^t(x) \left(1 - \frac{f_2(x)}{f_m}\right) \quad (25)$$

The foregoing identity is a fundamental statement about the structure of the two communities. From the definition of $I_M(a)$, $E_M(a)$ and the identify, we deduce

$$\sum_{x \in I_M(a)} u_M^t(x) > 1 - \frac{1 - \lambda_2^t}{a}, \quad \sum_{x \in E_M(a)} u_M^t(x) < \frac{1 - \lambda_2^t}{a} \quad (26)$$

with analogous inequalities for M replaced by m

We have seen that

$$u_2(x) = \frac{u_M^t(x) - u_m^t(x)}{\lambda_2^t(f_M - f_m)} + r(x), \quad (27)$$

where $r(x) = \lambda_2^t \times (\text{linear combination of } B_{xM}^t, B_{xm}^t)$. From the basic identity in Eq.(25), we deduce

$$\begin{aligned} \sum_{x \in I_M} u_2(x) &\geq \frac{1}{a[1 - (\lambda_1^t - \lambda_2^t)](f_M - f_m)} + \epsilon, \\ \sum_{x \in I_m} u_2(x) &\geq \frac{1}{a[1 - (\lambda_1^t - \lambda_2^t)](f_M - f_m)} + \epsilon', \end{aligned} \quad (28)$$

where ϵ, ϵ' are $o(1 - \lambda_2^t)$.

From Eq.(28), we notice that the gap of eigenvector value is $\sum_{x \in I_M} u_2(x) - \sum_{x \in I_m} u_2(x) \geq \frac{2}{a[1 - (\lambda_1^t - \lambda_2^t)](f_M - f_m)} + o(1 - \lambda_2^t)$ and its lower bound increases with the stability $\Theta(t) = \lambda_1^t - \lambda_2^t$. Thus, we declare that larger stability of communities will extend the eigenvector gap between them and thus enhance the significance of the community structure. Furthermore, one can easily extend the stochastic dynamic system to k -state, $k \geq 2$, in which the stability $\Theta(t) = \lambda_k - \lambda_{k+1}$ can also indicate the significance of structure owning k communities.

4. Estimate the spin correlation. – The spin correlation matrix C is very important for the Potts dynamic. In this section, we propose a novel way to calculate C using a new hierarchical block model method based on different granularity(resolution). Stochastic block model [9] is a useful tool to detect communities from networks or dynamical networks. However, the existing block model methods are restricted to the specific task of community detection and not suitable to models which need to extract multiple levels structure of hierarchical networks. In this part, the stochastic block model is extended to a multilevel form and exactly coincides with the dynamical process of the Potts model.

Let $A_{n \times n}$ be the adjacent matrix of network N , where n is the number of nodes. Suppose all nodes of N are divided into L ($1 \leq L \leq n$) blocks, denoted by $B_{n \times L}$, where $b_{il} = 1$ if node i is in block l , otherwise $b_{il} = 0$. When each block is considered to be inseparable, the granularity of network N can be measured by the number of blocks $g = L$. As g decreases from n to 1, the granularity of N degenerates from the finest to the coarsest. Let B_g denotes the block matrix B with a granularity g . In particular, we have $B_1 = I_{n \times n}$. Let matrix $Z_{L \times K}$ ($1 \leq K \leq L$) denotes such communities, where K is the community number and $z_{lk} = 1$ if block l is labeled by community k ,

otherwise $z_{lk} = 0$. Given Z , define $\Xi_{K \times n}$, where ξ_{kj} denotes the probability that any node out of community k expects to couple with node j ; and define $\Omega(\omega_1, \dots, \omega_K)^T$, where ω_k denotes the prior probability that a randomly selected node will belong to community k . It is easy to show that spin correlation matrix at level g is $C_g = B_g Z \Xi$ so calculating C_g is corresponding to estimate B_g , Z and Ξ .

Define $X = (K, Z, \Xi, \Omega)$ be a pattern unit of network N with respect to B_g . According to the principle of maximizing the posterior probability, the optimal X for a given network N under B_g will be one with the maximal posterior probability. Moreover, we have

$$P(X|N, B_g) \propto P(N|X, B_g)P(X|B_g) \quad (29)$$

where $P(X|N, B_g)$, $P(N|X, B_g)$, and $P(X|B_g)$ denote the posteriori of X given N and B_g , the likelihood of N given X and B_g , and the priori of X given B_g , respectively.

As discussed above, an optimal X will be the one with the maximal $P(X|N, B_g)$ and to maximize $P(X|N, B_g)$ is to maximize the product of $P(N|X, B_g)$ and $P(X|B_g)$. For a given K , the term $P(X|B_g)$ is a constant, and thus, to maximize $P(X|N, B_g)$ is to maximize $L(N|X, B_g)$.

Let $L(N|X, B_g) = \ln P(N|X, B_g)$, and we have

$$L(N|X, B_g) = \sum_{l=1}^L \sum_{b_{il} \neq 0} \sum_{k=1}^K \prod_{j=1}^n f(\xi_{kj}, a_{ij}) w_k \quad (30)$$

where $f(x, y) = x^y (1-x)^{1-y}$.

Considering the expectation of $L(N, Z|X, B_g)$ on Z , we have:

$$\begin{aligned} E[L(N, Z|X, B_g)] \\ = \sum_{l=1}^L \sum_{b_{il} \neq 0} \sum_{k=1}^K \gamma_{lk} (\sum_{j=1}^n (\ln f(\xi_{kj}, a_{ij})) + \ln \omega_k) \end{aligned} \quad (31)$$

where $E[z_{lk}] = \gamma_{lk} = P(y = k|b = l, X, B_g)$, i.e., the probability of block l will be labeled as community k given X and B_g . Let $J = E[L(N, Z|X, B_g)] + \lambda (\sum_{k=1}^K w_k = 1)$, we have:

$$\begin{cases} \frac{\partial J}{\partial \xi_{kj}} = 0 \\ \frac{\partial J}{\partial \omega_k} = 0 \\ \frac{\partial J}{\partial \lambda} = 0 \end{cases} \Rightarrow \begin{cases} \xi_{kj} = \frac{\sum_{l=1}^L \sum_{b_{il} \neq 0} a_{ij} \gamma_{lk}}{\sum_{l=1}^L \sum_{b_{il} \neq 0} \gamma_{lk}} \\ \omega_k = \frac{\sum_{l=1}^L \sum_{b_{il} \neq 0} \gamma_{lk}}{\sum_{k=1}^K \sum_{l=1}^L \sum_{b_{il} \neq 0} \gamma_{lk}} \\ = \frac{\sum_{l=1}^L \sum_{b_{il} \neq 0} \gamma_{lk}}{n} \end{cases} \quad (32)$$

Let $P(y = k|v = i)$ be the probability that node i belongs to community k given X and B_g , We have: $\gamma_{lk} = P(y = k|b = l, X, B_g) = \sum_{b_{il} \neq 0} \frac{1}{\sum_{i=1}^n b_{il}} P(y = k|v = i)$ where $\frac{1}{\sum_{i=1}^n b_{il}}$ is the probability of selecting node i from block l . According to the Bayesian theorem, we have:

$$P(y = k|v = i) = \frac{P(y = k)P(v = i|y = k)}{\sum_{k=1}^K P(y = k)P(v = i|y = k)}. \quad (33)$$

and

$$P(y = k)P(v = i|y = k) = \prod_{j=1}^n f(\xi_{kj}, a_{ij}) \omega_k \quad (34)$$

Thus

$$\gamma_{lk} = \frac{1}{\sum_{i=1}^n b_{il}} \times \sum_{b_{il} \neq 0} \frac{\prod_{j=1}^n f(\xi_{kj}, a_{ij}) w_k}{\sum_{k=1}^K \prod_{j=1}^n f(\xi_{kj}, a_{ij}) w_k} \quad (35)$$

As a conclusion, a local optimum of maximizing Eq.(29) will be guaranteed by recursively calculating Eq.(32) and Eq.(35) with granularity g . An optimal pattern unit $X = (K, Z, \Xi, \Omega)$ is calculated given B_g and consequently spin correlation matrix $C_g = B_g Z \Xi$.

For a given network, the hierarchical calculation process of C_g as g decreases from n to 1 can be incrementally proceeded as follows: First, constructing the ground layer by taking each node as one block, and $L = g = n$, $B_g = I_{n \times n}$. Thus, $C_n = B_g Z \Xi = A$. Then clustering it into $n - 1$ communities by selecting a model X_n with a maximum $P(X_n|N, B_n)$. Second, according to X_n , form B_{n-1} by capsuling each cluster in the ground layer as one block. $B_{n-1} = B_n \times Z_{n-1}$ and $C_{n-1} = B_{n-1} Z \Xi = A$. Then clustering these $n - 1$ blocks into $n - 2$ communities by calculating a new model X_{n-1} with a maximum $P(X_{n-1}|N, B_{n-1})$. Repeat the second step to construct more layers until the process converges, i.e., all blocks are grouped into only 1 communities. After calculating all n layers C_g , the average spin correlation matrix C can be taken as the hierarchical average $C = \langle C_g \rangle$, $g = 1, \dots, n$. One can easily find that the process coincides with the hierarchical dynamical process of Potts model described in Fig.1.

5. Experiments. – To show that the model can uncover hierarchical structures in different scales, Fig.2 and Fig.3 give two examples of the multi-level community structures, RB125 network [12] and H13-4 network [4]. In both examples, the most persistent Λ reveals the actual number of hierarchical levels hidden in a network. The significance of such levels can be quantified by their corresponding length of persistent time. Longer the time persists, more robust the configuration is. From Fig.2(b) and Fig.3(b), we can observe 25 and 16 are the optimal numbers of communities in RB125 and H13-4 networks owning the longest persistence, respectively. However, 5 modules and 4 modules are also reasonable partitions which show another fuzzy level of the hierarchical networks. These results are perfectly consistent with the generating mechanisms and hierarchical patterns of these two networks.

Furthermore, we also show that the variation tendency of stability $\Theta(\tau)$ in the two cases shed a light on the spin configuration. From Fig.2(b) and Fig.3(b), there are some local maximal values representing better community structure. Thus, we can find these local maximal timescales τ corresponding to the desirous number of communities and apply G^τ to a specific partition method. Furthermore, the stability will reach the lowest value at the end time of all Λ . The stability begins to increase when it transits to a new state. One can use $\Theta(\tau)$ to estimate the modularity property of complex networks, and the larger the Θ is, the stronger the network community structure. So, one can

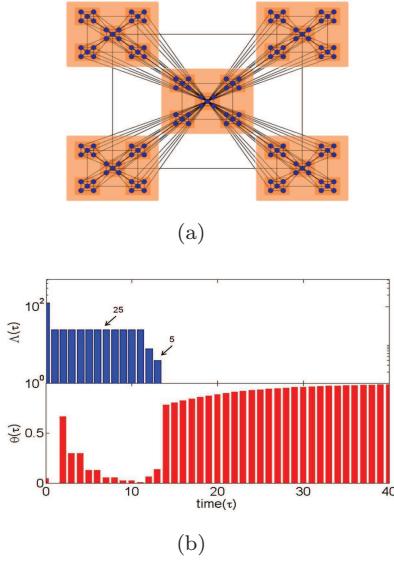


Fig. 2: (a) Structure of $RB125$, with 25 dense communities and 5 sparse communities, are highlighted in the original network. (b) The value of $\Lambda(\tau)$ and $\Theta(\tau)$ versus time τ .

find the largest corresponding Θ value for a specific number of community Λ and use it to indicate the robustness of modularity structure. For $H13-4$ shown in Fig.3(b), the stability of 16 communities structure, $\Gamma(16) = 0.62$ when $\tau = 3$, is larger than $\Gamma(4) = 0.43$ when $\tau = 12$. This indicates that the community structure containing 16 modules is more robust than community structure containing 4 modules. Similarly, for $RB125$ network shown in Fig.2(b), $\Gamma(25) = 0.71$ corresponding to 25 communities structure when $\tau = 3$ is larger than $\Gamma(5) = 0.18$ when $\tau = 13$. The robustness of community structure indicated by soft stability Θ favors finer but obvious modules which are reasonable for many real networks. In addition, the difference between the stability measure we proposed and the modularity Q [2] is emphasized. We also applied our framework to the hierarchical network with different modular sizes and some representative real networks. Finally, the relationships between our work and some famous concepts proposed in [6] and [8] are analyzed. These results are shown in the part of **Supplementary Material** [15].

6. Conclusion. — In summary, we have presented a more theoretically-based community detection framework which is able to uncover the connection between network's community structures and spectrum properties of Potts model's local uniform state. Important information related to community structures can be mined from a network's spectral significance through a Markov process computation, such as the stability of modularity structures and the optimal number of communities. Our method does not provide a unique optimal partition for the graph. Rather, we obtain number of stable levels and stability at each level over different layers of the hierarchical structure. Its effectiveness and efficiency have been demon-

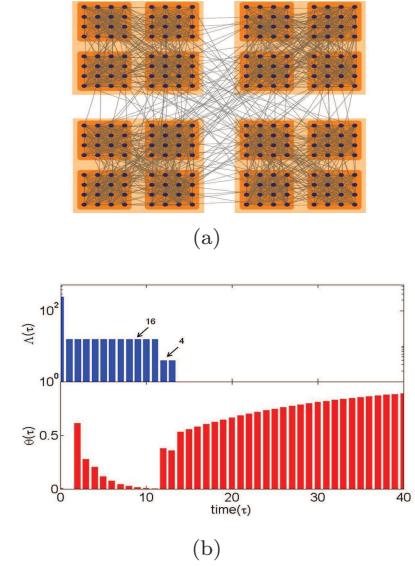


Fig. 3: (a) Structure of $H13-4$, with 16 dense communities and 4 sparse communities, are highlighted in the original network. (b) The value of $\Lambda(\tau)$ and $\Theta(\tau)$ versus time τ .

strated and verified both theoretically and experimentally.

The authors are separately supported by NSFC grants 11131009 and 71071090.

REFERENCES

- [1] NEWMAN.M.E.J and GIRVAN.M, *Phys. Rev. E*, **69** (2004) 026113.
- [2] NEWMAN.M.E.J, *Proc.Natl.Acad.Sci*, **103** (2006) 8577-8582.
- [3] FORTUNATO.S and BARTHELEMY.M, *Proc.Natl.Acad.Sci*, **104** (2007) 36.
- [4] ARENAS.A, FERNANDEZ.A, GOMEZ.S, *New. J. Phys.*, **10** (2008) 053039.
- [5] ARENAS.A, DIAZ-GUILERA.A, PEREZ-VICENTE.C.J, *Phys. Rev. Lett.*, **96** (2006) 114102.
- [6] DELVENNE.J.C, YALIRAKI.S.N and BARAHONA.M, *Proc.Natl.Acad.Sci*, **107(29)** (2010) 12755-12760.
- [7] BLATT.M, WISEMAN.S and DOMANY.E, *Phys. Rev. Lett.*, **76** (1996) 3251-3255.
- [8] REICHARDT.J and BORNHOLDT.S, *Phys. Rev. Lett.*, **93** (2004) 218701.
- [9] KARRER.B, NEWMAN.M.E.J, *Phys. Rev. E*, **83** (2011) 016107.
- [10] CAPOCCI.A, SERVEDIOA.Y.D.P, CALDARELLI.G, COLAIORI.F, *Physica A*, **352** (2005) 669-676.
- [11] ZHANG.X.S, WANG.R.S, WANG.Y, WANG.J, QIU.Y, WANG.L, CHEN.L, *Eur. Phys. Lett.*, **87** (2009) 38002.
- [12] RAVASZ.E and BARABÁSI.A.L, *Phys. Rev. E*, **67** (2003) 026112.
- [13] ZHOU.H, *Phys. Rev. E*, **67** (2003) 041908.
- [14] SHI.J AND MALIK.J, *IEEE Tans.On Pattern Analysis and Machine Intelligent*, **22(8)** (2000) 888-904.

[15] Please download the supplementary material file from the website
<http://doc.aporc.org/wiki/Hierarchical%20stability>